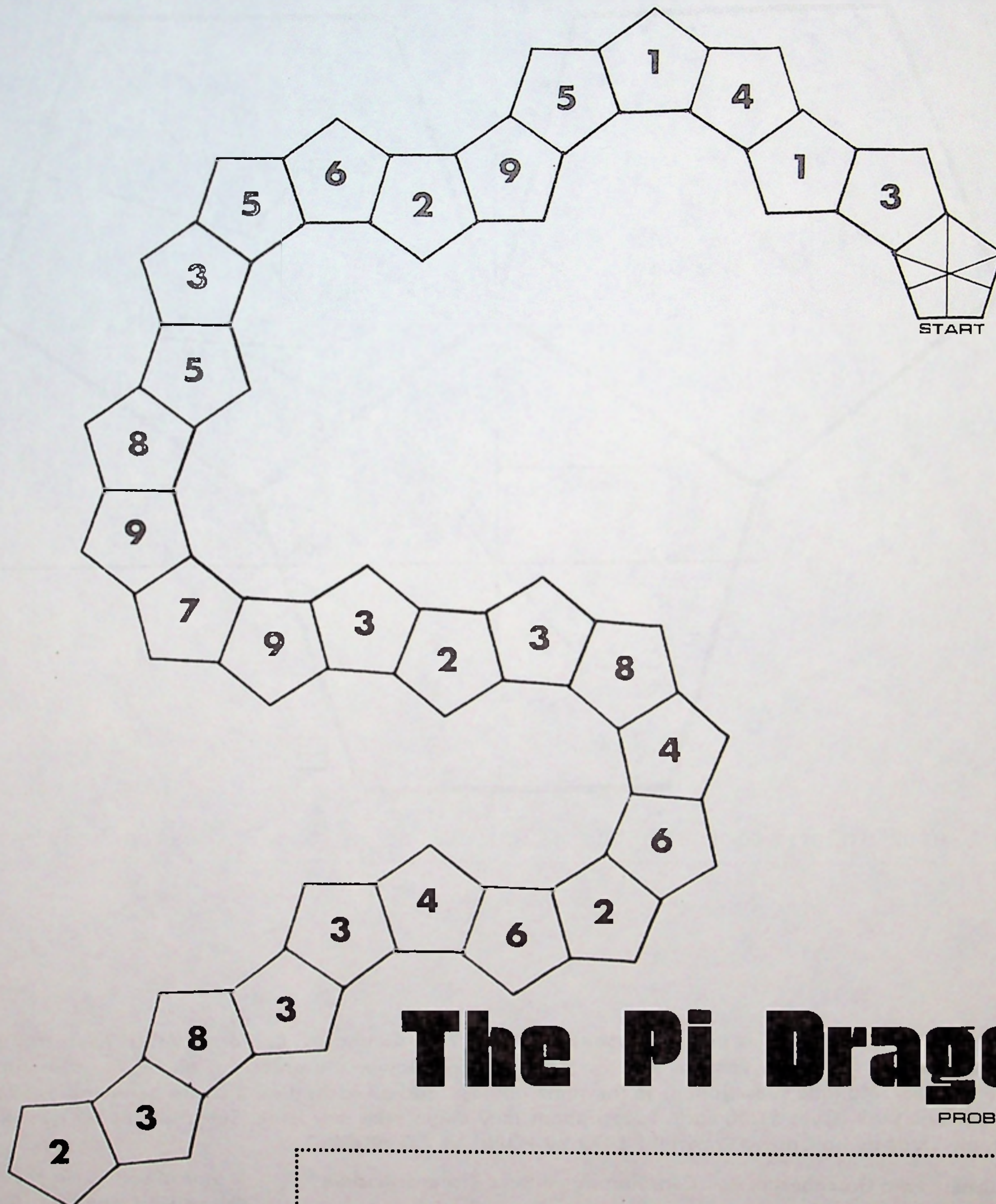
VOL 1 NO 6 SEPT 73

Popular Computing

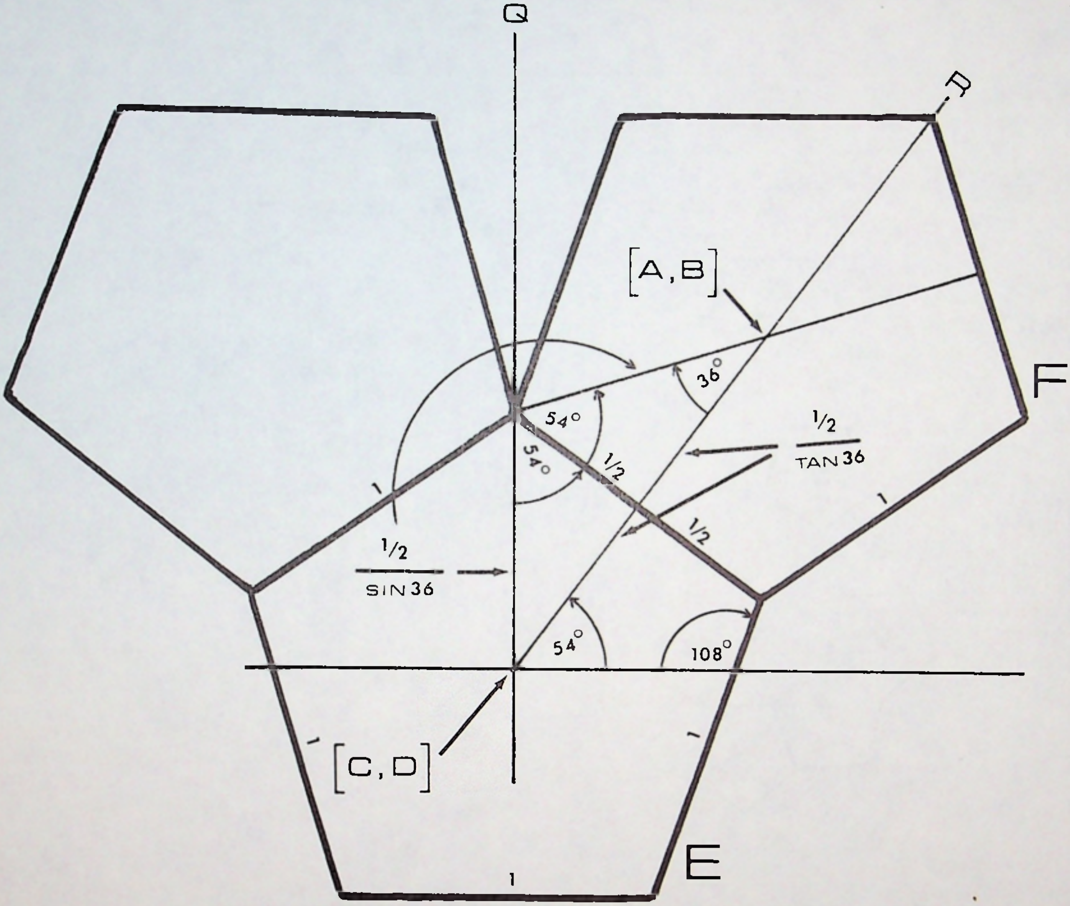
BOX 272 CALABASAS, CA 91302

6



The Pi Dragon

PROBLEM 12



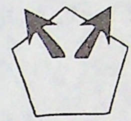
Useful information for the calculations of the Pi Dragon Problem.

POPULAR COMPUTING is published monthly at Box 272, Calabasas, California 91302. Subscription rate in the United States is $15 per year, or $12 if remittance accompanies the order. For all foreign subscriptions, add $4.50 per year. Multiple subscriptions to the same address, add $5 each; thus, 3 copies per month is $25 per year, U.S. delivery. Back issues $1.50 each. Subscriptions may begin with any issue. Subscriptions for qualified undergraduate students half price. Copyright 1973 by POPULAR COMPUTING.

```
314159265358979323
846264338327950288
[illegible]9716939937510582
097494459230781640
628620899862803482
534211706798214808
651328230664709384
460955058223172535
940812848111745028
410270193852110555
964462294895493038
196442881097566593
344612847564823378
678316527120190914
564856692346034861
045432664821339360
726024914127372458
700660631558817488
152092096282925409
171536436789259036
001133053054882046
652138414695194151
160943305727036575
959195309218611738
193261179310511854
807446237996274956
735188575272489122
793818301194912983
[illegible]7336244065664308
602139494639522473
719070217986094370
277053921717629317
675238467481846766
940513200056812714
526356082778577134
275778960917363717
872146844090122495
343014654958537105
079227968925892354
201995611212902196
086403441815981362
977477130996051870
721134999999837297
804995105973173281
609631859502445945
534690830264252230
825334468503526193
118817101000313783
875288658753320838
142061717766914730
359825349042875546
873115956286388235
378759375195778185
[illegible]8053217122680661
300192787661119590
9216420198
```

See the Figure on the cover page.

A regular pentagon forms a sort of arrow head:



A chain of pentagons is formed by appending a new pentagon to an old one on one of the two sides of the arrow head. The choice of side is determined by the odd or even nature of the digits in the decimal expansion of π. The first 29 links of the dragon are shown in the cover Figure.

Problem: Where is the center of the 1000th link of the dragon? Assume that the START pentagon is centered at the origin and that the pentagons have unit sides.

The drawing on page 2 shows the information needed to proceed from pentagon to pentagon of the Pi Dragon. At the stage where pentagon E is known (centered at C,D and pointing in direction Q), the calculation must determine the coordinates A,B of pentagon F, which points in direction R, 54° clockwise from Q. The case shown is for a turn to the right, as dictated by an even digit in π.

The 1000 digits of π shown at the left were keypunched 50 digits per card, in columns 1-50. The listing shows 55 lines of 18 digits and one line of 10 digits.

(1) Write a program to read the 20 cards and print the 56 lines. The program can be in assembly language, or any higher level language.

(2) Indicate, for any such program, how to change the parameter 18 to any other value.

(3) Indicate a test procedure for the program.

PROBLEM 13

THE DISTRIBUTION OF NUMBERS--COMPUTER THEORY

by R. W. Hamming

Bell Laboratories Murray Hill, New Jersey

It is reasonably well known that the fraction of physical constants whose mantissas (in scientific notation) are less than x is

$$D(x) = \frac{\log x}{\log b}$$

where b is the base of the number system, usually 10. Thus $D(x)$ is the probability of observing a mantissa which is less than x when picking a "random physical constant." A theoretical derivation for this observed cumulative probability distribution can be made based on the assumption that the distribution is not changed when the individual members are multiplied by some non-zero constant k; that is,

$$D(kx) = D(k) + D(x).$$

This same distribution arises when we examine the question of the distribution of floating point numbers in a computing machine. In particular, we shall show that the processes of multiplication and division tend to produce this distribution.

As a first step we consider a one decimal digit computer, and examine the 81 possible products rounded to one digit (of the products ending in 5 we round up half the time and down half the time, leaving the single product 5x5 to round up). From this experiment, we get the following table of products:

	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9
2	2	4	6	8	1	1	1	2	2
3	3	6	9	1	2	2	2	2	3
4	4	8	1	2	2	2	3	3	4
5	5	1	1	2	3	3	4	4	5
6	6	1	2	2	3	4	4	5	5
7	7	1	2	3	3	4	5	6	6
8	8	2	2	3	4	5	6	6	7
9	9	2	3	4	4	5	6	7	8

If we regard each of the original numbers as being equilikely, then the 81 products are all equilikely, and column 2 of the next table gives the corresponding frequencies, while column 3 gives the cumulative frequencies. Even in this crude model we can see how the distribution of the products is coming reasonably close to the distribution D(x).

	one digit		two digit		infinite digit		D(x)
digit	no.	cum.	obs. no. cum.	%	%	cum.	%
1	10	10	1871	23.1	24.1	21.4	30.1
2	18	28	3377	41.7	18.3	42.4	47.7
3	12	40	4565	56.4	14.5	56.9	60.2
4	12	52	5528	68.3	11.7	68.6	69.9
5	8	60	6309	78.0	9.5	78.1	77.8
6	9	69	6931	85.7	7.6	85.7	84.5
7	4	73	7431	91.9	6.1	91.8	90.3
8	5	78	7811	96.6	4.6	96.4	95.4
9	3	81	8100	100.1	3.4	99.8	100.0

Using a computer, we do the corresponding 8100 two decimal digit products to get the next two columns. Finally, using calculus, we get the infinite digit distribution in the next two columns. Comparing these results, we see that the two digit experiment gives results reasonably close to the infinite digit case so that two digit arithmetic is not too severe an approximation to actual practice. If we really doubted this, we could do the three digit case (using symmetry to save almost half the cases), but it does not seem worth it at this point. Finally, the last column gives the distribution D(x), and we see that starting with uniform distributions in x and y, the product x*y is already close to D(x).

Suppose that instead of the uniform distribution we took the factor x from the distribution D(x) and y from any other distribution you please. What would we get? To try it experimentally, we can take the midpoint number of each of the 90 intervals, 1.0 to 1.1, 1.1 to 1.2,

1.2 to 1.3,...,9.9 to 10.0 as a typical value for the numbers in their corresponding intervals. The probabilities of falling in these intervals are, of course,

$$\begin{array}{l} D(1.1) - D(1.0) \\ D(1.2) - D(1.1) \\ D(1.3) - D(1.2) \\ \cdots\cdots\cdots \\ \cdots\cdots\cdots \\ D(10.0) - D(9.9), \end{array}$$

for the x numbers, and the corresponding numbers for the y distribution whatever it is chosen to be. What will we get? For each y that you pick, the effect of multiplying those from the x distribution, because of the earlier result,

$$D(kx) = D(k) + D(x)$$

will merely reproduce the distribution D(x). Since this happens for each y value that you pick, regardless of the distribution of the y's, the result will be that the products will fall in the distribution D(x). Thus, if one factor comes from D(x) then the products will fall in the distribution D(x) regardless of what distribution the other factor comes from. This is called the "persistance of the distribution D(x)." If you try them, computer experiments will bear this out. A particularly spectacular illustration of this persistance is the example of forming a product of N factors

$$x_1 * x_2 * x_3 * \ldots x_N$$

If any one of the factors is from the distribution D(x), then the resulting product is also from the distribution.

We now turn to division. First we show that if x is from D(x) then so is 1/x. We set x*y b (the number base). Then each of the following steps is obvious, after a little thought. Let P(x) be the distribution of 1/x. We have the definition

$$D(x) = \text{Prob}\{t \leq x\}$$

Then

$$\text{Prob.}\left\{t \leq \frac{b}{x}\right\} = \text{Prob.}\left\{\frac{1}{t} \geq \frac{x}{b}\right\}$$

$$= 1 - \text{Prob.}\left\{\frac{1}{t} \leq \frac{x}{b}\right\}$$

$$= 1 - \text{Prob.}\left\{\frac{b}{t} \leq x\right\}$$

$$= 1 - D(b/x)$$

$$= 1 - \frac{\log(b/x)}{\log b} = \frac{\log x}{\log s}$$

$$= D(x)$$

Using this result we see that the behavior of mantissas in the process of division is reduced to the already solved case of their behavior in multiplication. Again, computer experiments, like those we did for multiplication, will verify this remark. When the original numbers x, y, and z, are all from the flat, equilikely distribution, the distribution of

$$x*y/z$$

is particularly close to the distribution D(x).

Let us summarize what we have learned. First, the distribution D(x) that first arose when considering the distribution of physical constants also arises from the processes of multiplication and division <u>provided</u> we assume some reasonable distributions for the initial numbers. Thus, we have an alternate basis for expecting to find the distribution D(x). Second, simple computer experiments enable us to see how numbers combine; indeed, in some ways these experiments teach us more than do the elaborate mathematical derivations, and in fact were the way the phenomenon was first discovered.

<u>EXERCISES</u>

1. Experimentally check that if x is taken from the distribution D(x), (meaning that the fraction of numbers less than x is (log x)/(log b), where by is the number base), then 1/x also has the same distribution.

2. Extend the results of table 1 to three decimal digit arithmetic, and summarize as in table 2.

3. Using the method of dividing the range into 90 subintervals and selecting the mid-value as being typical of those in the subinterval, examine the distribution of the 8100 products x*y and the 729,000 terms x*y/z assuming (a) each factor is selected from an equilikely distribution, (b) z is from D(z) and the others are from any other distributions you please.

Speaking of Languages...

BY ROBERT TEAGUE

[Beginning a monthly column of discussion of programming languages. Professor Teague is at California State University, Northridge.]

The objectives of this column are:

1) To have fun with the various programming languages.

2) To encourage more communication among those persons interested in programming languages.

3) To explore the facets and pitfalls of any and all programming languages.

4) To promote the use of ANSI standards by uncovering the quirks due to the variations among compilers.

Since this is the initial appearance of this column, it might be fun as well as educational to test our abilities in Algol, COBOL, Fortran, and PL/1 through a short quiz. The first person (determined by the earliest postmark) to submit correct answers to all questions will be declared the winner and his name will be published in a subsequent issue. Send your answers to:

Speaking of Languages...
POPULAR COMPUTING
Box 272
Calabasas, CA 91302

1. In PL/1, what does the following notation mean, and what data type does it imply for K and A?

$$K \rightarrow A$$

2. In COBOL, what level number must be assigned to ØVER-18?

```
IF ØVER-18 THEN GØ TØ DRAFT-STATUS.
```

3. In Algol, what kind of a division is implied by the following?

$$A \div B$$

4. In Fortran, is the following statement legal? What type is implied for K?

```
IF (K) GØ TØ 10
```

5. Recode the following Algol statement into <u>one</u> PL/1 <u>assignment</u> statement.

   L := <u>if</u> A > B <u>then</u> 4 <u>else</u> 0;

6. Recode the following Algol statement into <u>one</u> statement in Fortran, COBOL, and PL/1.

   <u>switch</u> A := L10, L20, L30, L40, L50;
   <u>goto</u> A [I] ;

7. Recode the following Algol <u>for</u> statements into <u>one</u> different statement in Algol, COBOL, Fortran, and PL/1.

   a) <u>for</u> J := M <u>step</u> 3 <u>until</u> 100 <u>do</u> (etc.)

   b) <u>for</u> L := 8, L+4 <u>while</u> L < 150 <u>do</u> (etc.)

8. Recode the following PL/1 assignment statement into <u>one</u> statement in Algol, COBOL, and Fortran.

```
A = B = 4;
```

9. Given the following PL/1 code, what will be the values of I and J after PL_1 is executed?

```
       DCL (A,B,C,D,E,I,J) FIXED BINARY;
       A = -8; B = 4;  C = 2;  D = 5;  E = 1;
       I = -A+B/C*(-D**2);
PL_1:  J = ¬(E & B < D);
```

10. Given the following COBOL entries, what will be the values possessed by E(1,2,1), E(2,1,2), D(1,1), D(2,2), C(2)?

```
01  A.
    02 FILLER PIC 99 VALUE 1.
    02 FILLER PIC 99 VALUE 2.
    02 FILLER PIC 99 VALUE 3.
    02 FILLER PIC 99 VALUE 4.
    02 FILLER PIC 99 VALUE 5.
    02 FILLER PIC 99 VALUE 6.
    02 FILLER PIC 99 VALUE 7.
    02 FILLER PIC 99 VALUE 8.
01  B REDEFINES A.
    02 C OCCURS 2.
        03 D OCCURS 2.
            04 E PIC 99 OCCURS 2.
```

The Raindrop Problem

PROBLEM 14

Given a square of side 1. Three uniformly distributed random numbers are drawn. The first two are uniform in the range zero to one, and are taken as the X and Y coordinates of a point in the unit square. The third number is uniform in the range zero to 1/2, and is taken as the radius of a circle. How many such triplets must be chosen and plotted so that the unit square is completely covered?

Flowcharting the procedure for selecting the required random numbers and describing (or plotting) the circles is relatively easy. The critical step is that of determining when the area covered by the circles also covers the complete square.

There are many excellent algorithms for generating random numbers with a computer. Indeed, the literature of the field has only one topic that occurs more frequently; namely, sorting techniques.

If one has only a pocket calculator at his command, and it functions only in floating decimal, many of the standard algorithms are not applicable. For casual use of random numbers, perhaps the old inner product method is again of interest. As a computing algorithm, this method has been thoroughly discredited, for reasons given below, but as a hand method (where the user can monitor the output by eye) it has some merit.

The method proceeds like this. Begin with two 4-digit numbers. Form their product, which will have 8 digits or less. Delete the low order two digits and retain the next four digits. Using this new number, and the one preceeding it, repeat the multiplication, obtaining the next 4-digit number in the sequence. For example:

1234
5678
0066
3747
2473
2663
5855
5918
6498 and so on.

As an algorithm to be used as a subroutine in a computer program, this scheme has two defects:

(1) The algorithm can degenerate to zeros at any time. The useful cycle length is thus unpredictable and is, of course, a function of the starting values. The cycle length to the point of zero degeneration is also usually relatively short. The degeneration point can be delayed by various tricks, but the method is inherently unstable. At least, however, the point at which the output goes to zero can be detected readily.

(2) The scheme can degenerate in another way. Suppose that it has produced 10,000 good 4-digit numbers, but then the 10001st and 10002nd numbers are the same as the 9001st and 9002nd numbers. From that point on, the same set of 1000 numbers will be generated over and over. This form of degeneration is difficult to detect, and vitiates the scheme as a suitable computer random number generator.

Despite these serious defects, the scheme (which is the oldest algorithm suggested as a random number generator) lends itself to hand calculation. The sample sequence started above, carried through 175 terms, develops the ten digits in this distribution:

0	70
1	68
2	63
3	56
4	75
5	67
6	87
7	67
8	84
9	63

Testing these observed frequencies against the theoretical (70 in each case) gives a chi-squared value of 11.8. With 9 degrees of freedom, the probability level is around .25, which is acceptable for the frequency test.

COMPUTIST: a specialist in computing. One who is skilled in or practices the art of computing. One versed in computation.

COMPUTERIST: a specialist in computers. One skilled and knowledgeable about the construction and design of computers.

Booklet Review

Widespread use of desk calculators has revived interest in a lost art. Old tricks are suddenly new tricks again. For example, on machines with sufficient accumulator capacity, sums of squares and cross products (of small numbers) can be accumulated in one operation. Suppose we have this data:

X	Y
2	3
4	5
6	7
8	9

If the first pair is entered as a single number, 20000003, and that number is squared and summed, and the same procedure is followed for each pair, then the final result shows

12000002800000164

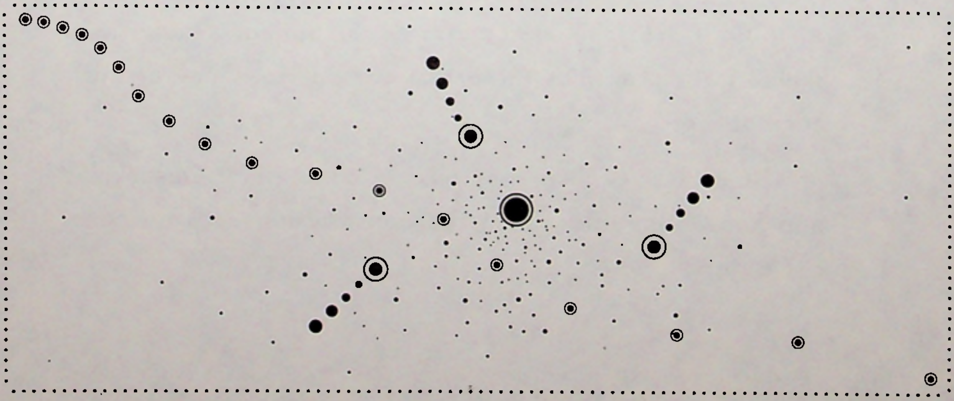
which is both sums of squares and twice the sum of the cross products. If this were done on an old rotary calculator, the secondary dial also shows

200000024

which gives the two simple sums. At present, this trick is not feasible on the electronic machines.

On the other hand, elaborate tables that were built to facilitate desk calculator work are again valuable. One such table is

Tables in Brief
Mathematical Tables for Use with Desk
    Calculators
Wayne White, June, 1951
RM-660
The RAND Corporation
1700 Main Street, Santa Monica, CA 90406
$2.00, 34 pages



The booklet contains tables (most of them on one page each) of common logarithms, natural logarithms, exponentials, 10^x, sine, cosine, tangent (and their inverse functions) in both degrees and radians, hyperbolic functions, square roots, gamma function, elliptic integrals, and Bessel functions--and all of these arranged for ease of interpolation, both linear and quadratic. For example, the following are entries in the logarithm table:

X	Log	A	B
2.25	.35218252	.00193020	-.00000428
2.26	.35410844	.00192166	-.00000424

To acquire log 2.2573, one calculates (for linear interpolation)

$$\log(2.25+.0073) = \log(2.25) + .73(.00193020 - .00000428)$$
$$= .3535884416$$
$$(.3535893 \text{ true})$$

For quadratic interpolation:

$$\log(2.2573) = \log(2.25) + AP + BP^2$$

where A and B are given in the table and P is the fraction .73.

$$= .3535892852$$

6

Log 6	0.77815125038364363250876679797960833596831874565280
Ln 6	1.79175946922805500081247735838070227272299069218300
$\sqrt{6}$	2.44948974278317809819728407470589139196594748065667
$\sqrt[3]{6}$	1.81712059283213965889121175632726050242821046314122
$\sqrt[5]{6}$	1.43096908110525550104522441314311690497264993966128
$\sqrt[7]{6}$	1.29170834209074660682750975536547798190706873345974
$\sqrt[10]{6}$	1.19623119885131548973381914341377367153754991741159
$\sqrt[100]{6}$	1.01807907781330729223910006217712625613813808471496
e^6	403.428793492735122608387180543388279605899897357129
π^6	961.389193575304437030219443652419898867217528081047
$\tan^{-1} 6$	1.40564764938026978095219340199580798810019803922253
6^{100}	6533186235000709060966902671580578205371437104729548 71543071966369497141477376
6^{1000}	1416610262383486172379625252491522441664047183091019 1322323547432140618947596486436347661333869287260068 9079493020294849159424026812116206945980466178442955 1222079310331298054959153716095905302794062411759800 34175030157226974281761555600362263128567590299511776 6865928620743763282329903251012486801237769145764828 1509578456812298622189041183773757009886461334209097 2756469661488216176894465388028416768338495326989675 1180872227673845961113513049578690252738029782817837 3192996646821057922983006955669892893734250898834079 2335737744719376598506908977135291983117722648269177 9471546576975170749934415155268398870734001917974451 5376022169572326825500613404406250310071013420041460 7696976757837002911389023284338696251543694980946202 137938610119300450795091488653253649628649410789376

N-SERIES